BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 8: e142847
doi: 10.3897/biss.8.142847

OPEN ACCESS

Conference Abstract

Deciphering Japanese Labels on Old Natural History Specimens

Hirochika Ueda‡, Yurika Saito‡, Haruo Tanaka‡, Keiichi Kaneko‡

‡ Tokyo University of Agriculture and Technology, Tokyo, Japan

Corresponding author: Hirochika Ueda (uedahirochika@gmail.com)

Received: 27 Nov 2024 | Published: 27 Nov 2024

Citation: Ueda H, Saito Y, Tanaka H, Kaneko K (2024) Deciphering Japanese Labels on Old Natural History Specimens. Biodiversity Information Science and Standards 8: e142847. https://doi.org/10.3897/biss.8.142847

Abstract

For natural history specimens, metadata such as collection locality, date, and collector's name are essential information ensuring the specimens' scientific value. However, specimens collected decades ago may have missing or illegible metadata.

The Takashi Kurobe's Rock and Mineral Collection*[1] consists of 883 rock, mineral, and fossil specimens stored in the Soil Science Laboratory in the Faculty of Agriculture at the Tokyo University of Agriculture and Technology. Of these, 611 specimens were collected through purchase, while the remaining 272 were collected by faculty members, students, and university associates. The specimens collected by faculty members are mainly accompanied by hand-written labels, and many of them contain partial information or are difficult to read because of the cursive scripts.

The Japanese language is represented by 48 Hiragana and Katakana characters each, and more than 10,000 Kanji characters. In addition, the labels of old natural history specimens were written in English and Romanized Japanese using the alphabet. Such a variety of characters makes it difficult to decipher handwritten labels on the specimens even for native Japanese speakers. This study reports on how label information in the Takashi Kurobe's Rock and Mineral Collection was deciphered to recover the scientific value of its specimens.

Following this study, the Takashi Kurobe's Rock and Mineral Collection*[1] has been made available as a digital archive with the photographs and metadata of each specimen on

© Ueda H et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

the web page of Nature and Science Museum, Tokyo University of Agriculture and Technology.

CASE 1

In Takashi Kurobe's Rock and Mineral Collection, there was a rock specimen with a single label with five characters in cursive scripts (Fig. 1). It was clear from their number and shapes that these five characters likely represent the location where the specimen was collected, but it was illegible due to the cursive scripts. The Center for Open Data in the Humanities is developing an application "miwo" that uses AI to recognize cursive scripts in documents. However, the accuracy of "miwo" is not yet high enough. Due to the above situation, deciphering the labels of this specimen was a dead end. Therefore, we decided to look not only at the label but also at the specimen itself.

Figure 1.

A volcanic rock specimen*[2] with a label written in illegible cursive scripts in Takashi Kurobe's Rock and Mineral Collection. These photos are licensed under Creative Commons Attribution-Noncommercial-Sharealike 4.0 International by Nature and Science Museum, Tokyo University of Agriculture and Technology.

This specimen was a dark green volcanic rock with uneven pores throughout. On the inner surface of some of the bubbles, fine yellowish-gray crystals were found. These features are common in boninite. In Japan, boninite is found only around the Bonin Islands, far south of Tokyo. Looking at the specimen with the bias that this specimen would be from the Bonin Islands, the idea that the first two characters on the labels were possibly "兄島 Ani-Jima" came to the mind of the authors. "Ani-Jima" is Ani Island, located in the northern part of the Bonin Islands. Furthermore, the place named three Kanji where the Boninites outcrop on Ani Island was searched, leading to the discovery of a place called "瀧の浦 Taki-No-Ura". As a result, the specimen's labels were deciphered as reading "Ani-Jima Taki-No-Ura", which is the Taki-No-Ura beach on Ani Island, in the Bonin Islands chain. This case clearly shows that the specimen itself is the largest source of the information in deciphering the label.

CASE 2

The Takashi Kurobe's Rock and Mineral Collection contains 21 limestone specimens with partial locality information (Fig. 2). However, their labels only gave local site-specific names without a broader geographic context. After further investigating the collections, seven limestone specimens whose labels included these local names with a broader area's name, such as Prefectural or city names, were found. This led to the conclusion that the 21 limestone specimens were from Mt. Kinsho, Ogaki City, Gifu Prefecture, Japan. Although the 21 limestone specimens were found to be from Mt. Kinsho, the detailed place names on the labels were not found in the present day. One of the 21 limestones was found with labels bearing the initials "T. W." The same initials were found on the specimens of the Takashi Kurobe's Rock and Mineral Collection, and it was assumed that these initials represented Tetsugoro Wakimizu, who was a professor of Geology and Soil Science at Tokyo Imperial University. Wakimizu published papers on the Akasaka Limestone in 1901–1902, which included a geological map of Mt. Kinsho, containing detailed place names on the labels mentioned above. According to this, the specimens' detailed localities were determined. The reason why these place names are not found today is that the geography has changed due to the limestone mining and these places have either changed or disappeared. This case shows that by comparing the fragmentary information on the labels with the previous articles of the time that specimens were collected, has supplemented the missing information and restored the scientific value of the specimens.

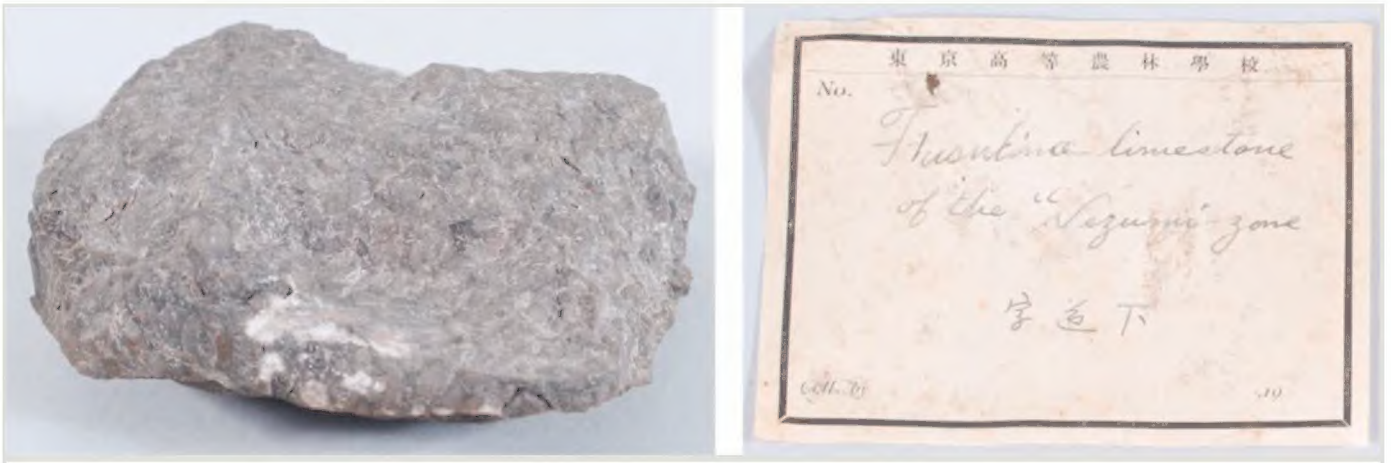

Figure 2.

One of the 21 limestone specimens*[3] with partial locality information on their labels in Kurobe's Rock and Mineral Collection. These photos are licensed under Creative Commons Attribution-Noncommercial-Sharealike 4.0 International by Nature and Science Museum, Tokyo University of Agriculture and Technology.

In this study, we went back to the basics to decipher specimen information while utilizing AI and a database of cursive Kanji. The specimens themselves are the most important source of information, and the specimen labels can be supplemented with information that is otherwise illegible or missing from the labels by the specimen's localities and collectors. Namely, once the specimen's localities and collectors are known, further information can be supplemented by examining the publications of the periods when the specimen was collected, thereby greatly restoring the specimen's scientific value.

Although the development of AI is remarkable, at present, we would continue to rely on both the experiences and knowledge of curators to investigate and sort the natural history collections.

Keywords

rock, mineral, fossil, university museum, cursive script

Presenting author

Hirochika Ueda

Presented at

SPNHC-TDWG 2024

Acknowledgements

We thank Professor Yasumoto Nakazawa and the museum staff for supporting our research, and Sakiko Ueda and the curator course students from 2017, 2018, and 2023, for helping our digital archiving project of Takahashi Kurobe's Rock and Mineral Collection.

Conflicts of interest

The authors have declared that no competing interests exist.

Endnotes

*1 https://archives.tuat-museum.org/s/agriculture/page/rockmineral

*2 https://archives.tuat-museum.org/sunaba_202209_dup/s/agriculture/item-link/2022-C4-013#?c=&m=&s=&cv=&xywh=-596%2C0%2C5127%2C2623

*3 https://archives.tuat-museum.org/sunaba_202209_dup/s/agriculture/item-link/2022-C1-008#?c=&m=&s=&cv=&xywh=-597%2C0%2C5065%2C2591